RS

#2

Europäisches Patentamt | European Patent Office | Office européen des brevets

jc921 U.S. PTO
09/838406
04/19/01

Bescheinigung | Certificate | Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. | Patent application No. | Demande de brevet n°

00480053.8

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

I.L.C. HATTEN-HECKMAN

DEN HAAG, DEN
THE HAGUE, 14/09/00
LA HAYE, LE

EPA/EPO/OEB Form 1014 - 02.91

Europäisches Patentamt | European Patent Office | Office européen des brevets

Blatt 2 der Bescheinigung
Sheet 2 of the certificate
Page 2 de l'attestation

Anmeldung Nr.:
Application no.: 00480053.8
Demande n°:

Anmeldetag:
Date of filing: 20/06/00
Date de dépôt:

Anmelder:
Applicant(s):
Demandeur(s):
International Business Machines Corporation
Armonk, N.Y. 10504
UNITED STATES OF AMERICA

Bezeichnung der Erfindung:
Title of the invention:
Titre de l'invention:
IP data transmission network using a route selection based on the level 4/5 information

In Anspruch genommene Prioriät(en) / Priority(ies) claimed / Priorité(s) revendiquée(s)

Staat:
State:
Pays:

Tag:
Date:
Date:

Aktenzeichen:
File no.
Numéro de dépôt:

Internationale Patentklassifikation:
International Patent classification:
Classification internationale des brevets:
/

Am Anmeldetag benannte Vertragstaaten:
Contracting states designated at date of filing: AT/BE/CH/CY/DE/DK/ES/FI/FR/GB/GR/IE/IT/LI/LU/MC/NL/PT/SE
Etats contractants désignés lors du depôt:

Bemerkungen:
Remarks:
Remarques:

IP DATA TRANSMISSION NETWORK USING A ROUTE SELECTION BASED ON THE LEVEL 4/5 INFORMATION

Technical field

The invention relates to IP data transmission network wherein the route is determined in each router of the data path by using a combination of metrics and relates in particular to an IP data transmission network using a route selection based on the level 4/5 information.

Background

When data packets are transmitted from a source workstation to a destination workstation in an IP data transmission network, the packets are routed from node to node in the network by a routing mechanism implemented by a router in each node of the data path.

Each IP datagram received by a node and which specifies a destination address other than the local node address is subject to the IP routing algorithm by the router of the node

which selects the next node for the datagram. For this, the router uses a routing table which contains information about the other routers of its own network and about IP networks to which the own network is attached.

The routing mechanism enables the optimal routing path to be determined. Such a path determination is based on a variety of metrics or a combination of metrics, such metrics being values resulting from algorithmic computations on a particular variable or values directly input by the router administrator. The comparison of the metrics allows the router to determine the optimal routes and therefore to establish the routing table.

Many different metrics have been used in routing algorithms. Sophisticated routing algorithms can base the route selection on multiple metrics, combining them in a manner resulting in a single metric. For this, several metrics may be used. Path length is the most common metric. Some routing protocols allow the network administrator to assign arbitrary costs to each network link. In this case, the path length is the sum of the costs associated with each link traversed. Another important metric to be used is the communication cost insofar as some companies may not care about performance whereas they care about operating expenditures. But other metrics may be used such as reliability (usually described in terms of bit error rate) of each network link, delay (the time required to move a packet from a source to a destination), bandwidth (available traffic capacity of a link) or load (the degree to which a network resource such a router is busy).

All the information used by the routing protocol to determine the metrics are coming from the layers up to layer 3. It must be reminded that in the internet protocols as well as in the Open System Interconnect (OSI) model, the layer which defines the packet delivery including routing is the third layer. But

the upper layers defining the application such as the Transmission Control Protocol (TCP) or the User Datagram Protocol (UDP) which constitutes the layer 4/5 in the network procedure are not used to determine metrics and therefore to build the routing table.

There have been several tentatives to use the level 4/5 information in the routing algorithm. All these approaches were based on assuming that the source and destination workstations provide information to the routers and support a specific protocol to set the inputs. These tentatives which were putting effort on the level 4/5 (for the TCP/IP model as well as for the OSI model) never got to stable and durable implementation and have been given up. Thus, a tentative has been implemented in OSPF using the configuration parameters called Type of Service (TOS). This mechanism requires the application to set the TOS field in all IP data generated by the source workstation. But the support of TOS in routing requires modifications in the application. Today, no application has been implemented in such a way. Therefore, the use of the TOS field has been dropped in recent OSPF implementations according to the latest RFC recommendations. Some other tentatives use a priority queue in the routers : the router tries to look at some field significant of level 4/5 application layer when receiving the IP data packets in order to set some priority in input/output buffer. Such a mechanism acts on the data transmission speed, but does not modify the routing path.

Summary of the invention

Accordingly, the main object of the invention is to provide a route selection in the routing algorithm of an IP data transmission network which depends upon the type of application used in a communication between a source workstation and a destination workstation.

Another object of the invention is to select a routing path in OSPF routing protocol according to the TCP or UDP port required in each router of a transmission path in an IP data transmission network.

The invention relates therefore to a data transmission system for transmitting packets of data from a source workstation to a destination workstation wherein the packets of data are transmitted over at least an IP network between an ingress node connected to the source workstation and an egress node connected to the destination workstation, wherein each router within the intermediary nodes along the data path from the ingress node to the egress node determines the best route in a routing table defined by the contents of a field contained in each packet of data being received. For this, the router of the ingress node comprises a configuration table which defines the contents of the field, the TOS field in the preferred embodiment, in function of information associated with a protocol having a level upper than the IP level, generally the 4/5 level such as TCP or UDP.

Brief description of the drawings

The above and other objects, features and advantages of the invention will be better understood by reading the following more particular description of the invention in conjunction with the accompanying drawings wherein :

-Fig. 1 is a block-diagram representing a data transmission network wherein the invention can be implemented.

-Fig. 2 is a schematic representation of an IP datagram with the various fields of the IP header.

-Fig. 3 shows the contents of the service type field of the IP datagram including the TOS bits.

-Fig. 4 is a schematic representation of the IP datagram showing the TCP/UDP header after the IP header.

-Fig. 5 is a flow-chart of the different steps which are implemented in an ingress node according to the invention.

-Fig. 6 is a schematic representation of configuration table used in the data transmission system according to the invention.

-Fig. 7 is a block-diagram representing the router of the ingress node according to the invention.

Detailed description of the invention

In reference to Fig. 1, a data transmission system according to the invention can include a source workstation 10 attached to a LAN 12 and which may access to an IP network 14 through an ingress node 16 in order to transmit packets of data to a destination workstation 18 connected to the network 14 by an egress node 20.

The optimal route which is determined by computing a combination of metrics, is defined in ingress node 20 according to the level 4/5 information as described below. Thus, such a route could be through intermediary node 22 with a first application whereas the optimal route could be through intermediary node 22 with a second application.

As an example, it can be assumed that the first application is a voice session VoIP requiring low delay but not requiring a large bandwidth and the second application is a data batch transfer in File transfer Protocol (FTP) without delay requirement but large throughput. Accordingly, the delay through intermediate node 22 is better than delay though intermediate node 24 but the overall bandwidth on this route is limited. The delay through intermediate node 24 is longer but a lot of free bandwidth is available on this second route. It must be noted that with shorter delay on the route through node 22, both sessions, voice and data, would be established on this route when using legacy OSPF, whereas the invention

based on the 4/5 level information enables two optimal routes to be established thereby improving the load balancing between all routes in the network.

As illustrated in Fig. 2, an IP datagram data includes an IP header containing the information necessary to send the packet in correct form over the route such as the source IP address or the IP destination address, and the IP data field 26. The header 24 includes a service type field 28 which is illustrated in Fig. 3. Such a service type field includes two bits for precedence (measure of the priority of the datagram), a MBZ bit (must be zero) reserved for future use and the Type Of Service bits 3-6 which characterize some parameters of the application such as the propagation time, the throughput and the reliability.

In an IP datagram, further to the IP header 24, the IP data field 26 is composed of a header TCP/UDP and data as illustrated in Fig. 4. Note that in both TCP and UDP protocols and other 4/5 level protocols, the header includes a field containing the source port number and a field containing the destination port number.

In the ingress node 16 (see Fig. 1), the sequence of steps represented in Fig. 5 are as follows when a frame is received (step 30). First, the protocol identification is made by extracting (step 32) the associated field in the IP header of the IP datagram (see Fig. 2). The source and/or destination port number is also extracted (step 34) from the 4/5 level header. Then, there is a lookup of a configuration table represented in Fig. 6 (step 36). Such a configuration table provides the TOS field corresponding to a given protocol and a given port number (source and/or destination). It must be noted that, insofar as there can be a lot of port numbers for each protocol being identified, such a table may not contain all possible cases.

The lookup of the configuration table determines whether an entry of the table is associated with the protocol being identified (step 38). If so, it is then determined whether the port number extracted from the 4/5 level header is identified in the configuration table (step 40). If it is the case, the TOS bits in the IP datagram are replaced (step 40) by the bits which are found in the configuration table. It must be noted that the TOS bits are previously all zeros corresponding to the default route. Finally, the routing task is achieved by the OSPF protocol or an equivalent protocol (step 44) before transmitting the frame over the network.

When the protocol being identified in the IP datagram does not correspond to any entry in the configuration table or when there is no port number for this protocol in the configuration table corresponding to the port number identified in the 4/5 level header, the TOS bits are set to zero or not changed if they were already set to zero (step 46). Then, the OSPF (or equivalent protocol) routing task is achieved by the router (step 44). Note that, in the latter case, the "all zeros" TOS defines only the default route.

As illustrated in Fig. 7, the router of ingress node 20 (see Fig. 1) includes a protocol processing unit 50 which identifies the protocol associated with the application in the received frame. Thus, such a protocol can be UDP plus Real Time Protocol (RTP) for the voice flow or TCP plus FTP for the data flow. Note that the protocol processing unit 50 looks for also the port number associated with the protocol in the frame header as already mentioned.

When the protocol and the port number are identified by protocol processing unit 50, the frame together with this information are transmitted to forwarding processing unit 52. The protocol and the port number enable the forwarding

processing unit 52 to look in configuration table 54 for determining the value of TOS bits corresponding to this protocol and this port number. Then, the forwarding processing unit 52 replaces in the frame the previous TOS bits (generally zero bits) by the value determined in configuration table 54 and stores the frame into frame buffer 56. The TOS value also enables the forwarding processing unit 52 to select an appropriate routing table identifying the route to be used, for example routing table 58 containing routing information for the UDP protocol or routing table 60 containing routing information for the TCP protocol. Then, forwarding processing unit 52 utilizes this routing information to select an output queue for transmitting the frame over the network. In the previous example, the voice frame using UDP will have a routing information that will enable the transfer of the voice frame from frame buffer 56 to output queue 62 for transmission to intermediate node 22, whereas the data frame using TCP will have a routing information that will enable the transfer of the data frame from frame buffer 56 to output queue 64 for transmission to intermediate node 24.

The routers of all the other nodes of the route such as intermediate node 22 or 24 include the same components except the configuration table since they have no need to determine and to change the value of the TOS field in the frame. In the same way as in the ingress node, the TOS value enables each router to select the appropriate routing table in order to know the routing information and to transmit the frame to the next node of the route.

CLAIMS

1. Data transmission system for transmitting packets of data from a source workstation (10) to a destination workstation (18) wherein said packets of data are transmitted over at least an IP network (14) between an ingress node (16) connected to said source workstation and an egress node (20) connected to said destination workstation, wherein each router within the intermediary nodes along the data path from said ingress node to said egress node determines the best route in a routing table defined by the contents of a field contained in each packet of data being received ;

said system being characterized in that the router of said ingress node comprises a configuration table (54) which defines the contents of said field in function of information associated with a protocol having a level upper than the IP level.

2. Data transmission system according to claim 1, wherein said field defined by said configuration table (54) determines the contents of the Type Of Service (TOS) field to be included in the IP header of said packet of data being received.

3. Data transmission system according to claim 2, wherein said field defined by said configuration table (54) is determined in the latter by the protocol identification and by the source and/or destination port number.

4. Data transmission system according to claim 1, 2 or 3 wherein, each router within the intermediary nodes (22 or 24) along the data path from said ingress node (16) to said egress node (20) includes a plurality of routing

tables (58, 60) corresponding each to the different contents of said field.

5. Process of transmitting packets of data in an IP network from a source workstation (10) to a destination workstation (18) between an ingress node (16) connected to said source workstation and an egress node connected to said destination workstation wherein each router within the intermediary nodes along the data path from said ingress node to said egress node determines the best route in a routing table; said process being characterized in that it consists in :

- extracting from each packet of data received in said router information associated with a protocol having a level upper than the IP level, and
- using said information for determining the best route by selecting one of a plurality of routing tables.

6. Process according to claim 5, wherein said information associated with a protocol having a level upper than IP level is the protocol identification and the source and/or destination port number.

7. Process according to claim 6, wherein said step of using said information consists in determining the value of a field in said packet of data which depends on said protocol identification and said source and/or destination port number.

8. Process according to claim 7 , wherein said field is the Type of Service (TOS) field in the IP header of said packet of data.

9. Process according to claim 8, wherein the value of said TOS field is modified in accordance with said protocol

identification and said source and/or destination port number in the router of said ingress node.

10. System comprising means for implementing the steps of the process according to any one of claims 5 to 9.

IP DATA TRANSMISSION NETWORK USING A ROUTE SELECTION BASED ON THE LEVEL 4/5 INFORMATION

Abstract

Data transmission system for transmitting packets of data from a source workstation (10) to a destination workstation (18) wherein the packets of data are transmitted over at least an IP network (14) between an ingress node (16) connected to the source workstation and an egress node (20) connected to the destination workstation, wherein each router within the intermediary nodes along the data path from the ingress node to the egress node determines the best route in a routing table defined by the contents of a field contained in each packet of data being received. For this, the router of the ingress node comprises a configuration table which defines the contents of the TOS field in function of information associated with the 4/5 level protocol such as TCP or UDP.

FIG. 1

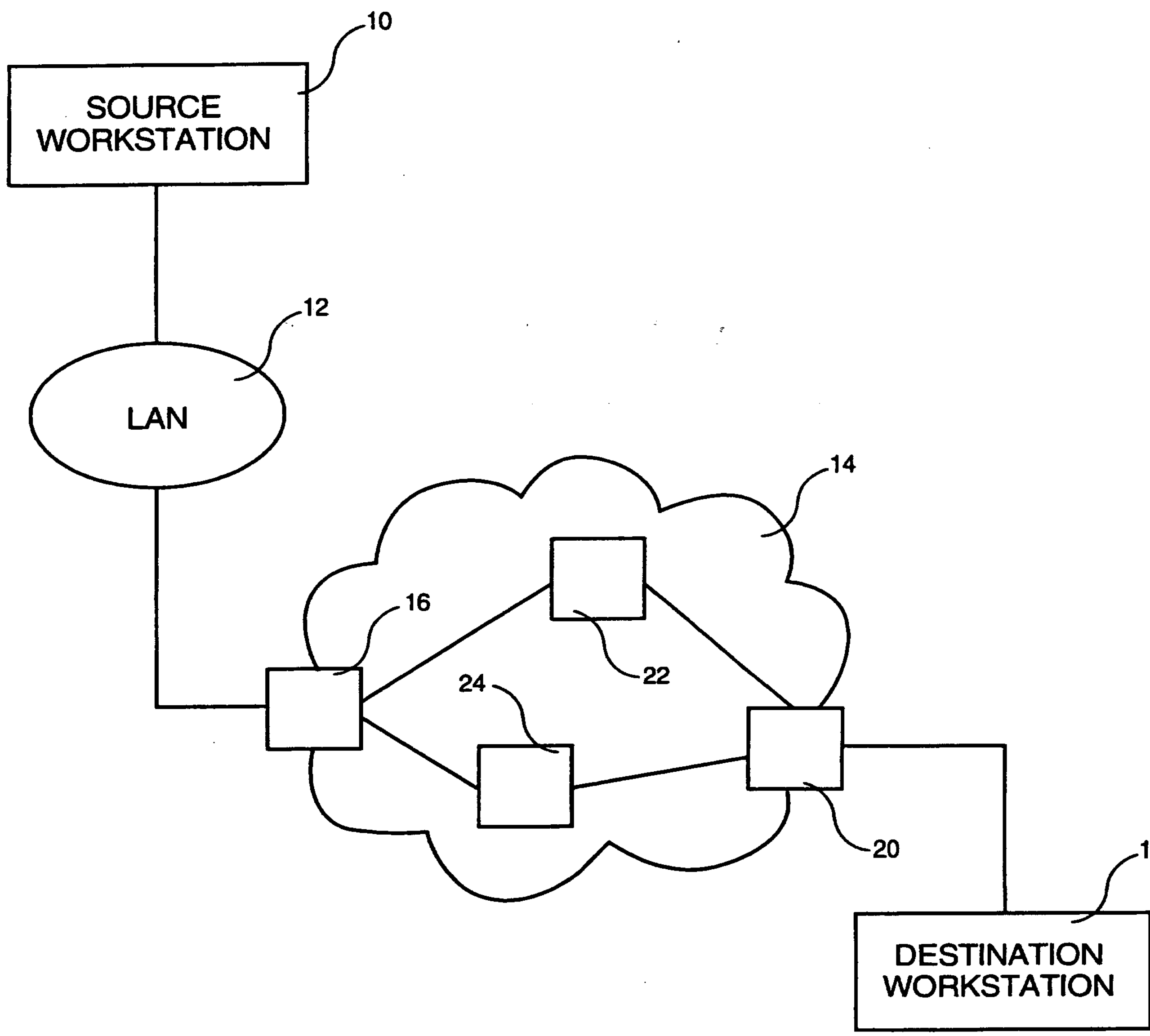
10
SOURCE WORKSTATION
12
LAN
14
16
22
24
20
18
DESTINATION WORKSTATION

FIG. 1

FR 9 00 008
Benayoun et al
2/4

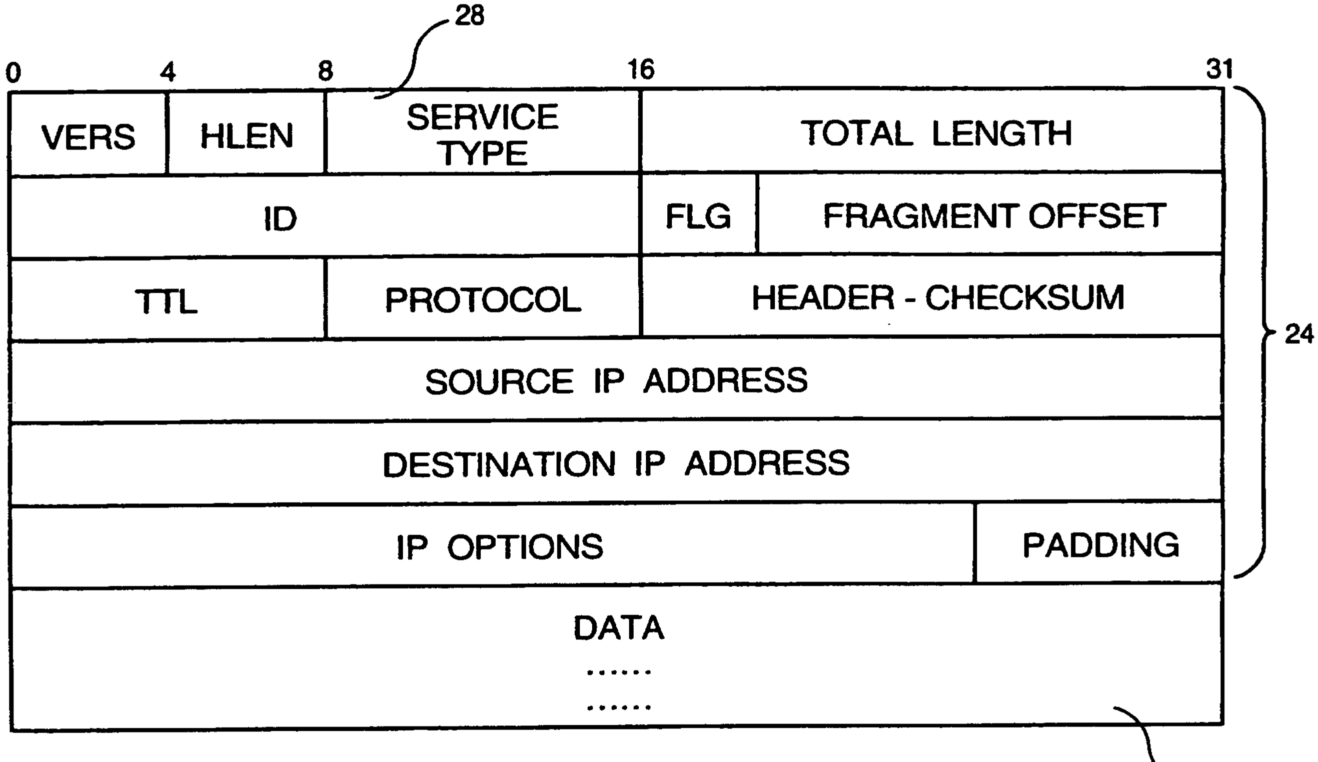

FIG. 2

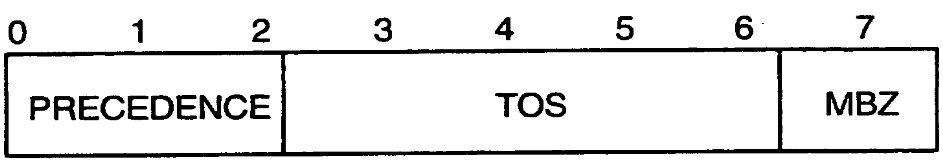

FIG. 3

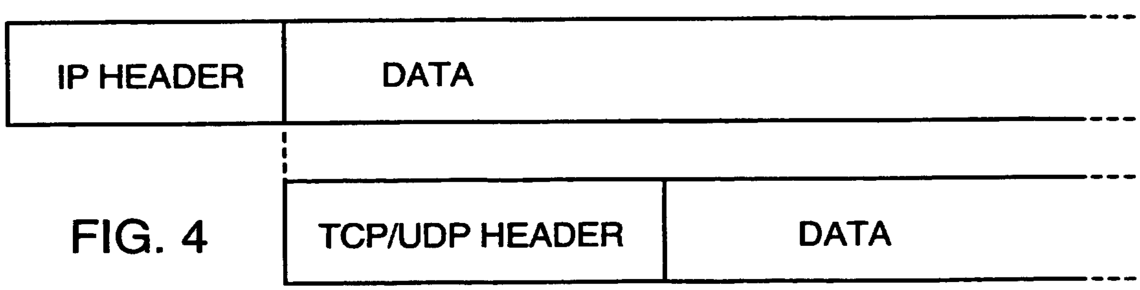

FIG. 4

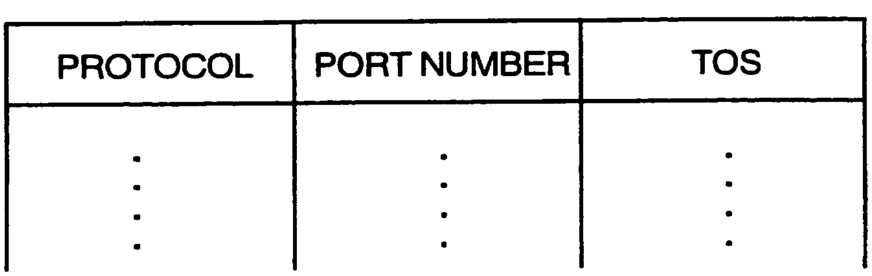

FIG. 6

FR 9 00 008
Benayoun et al
3/4

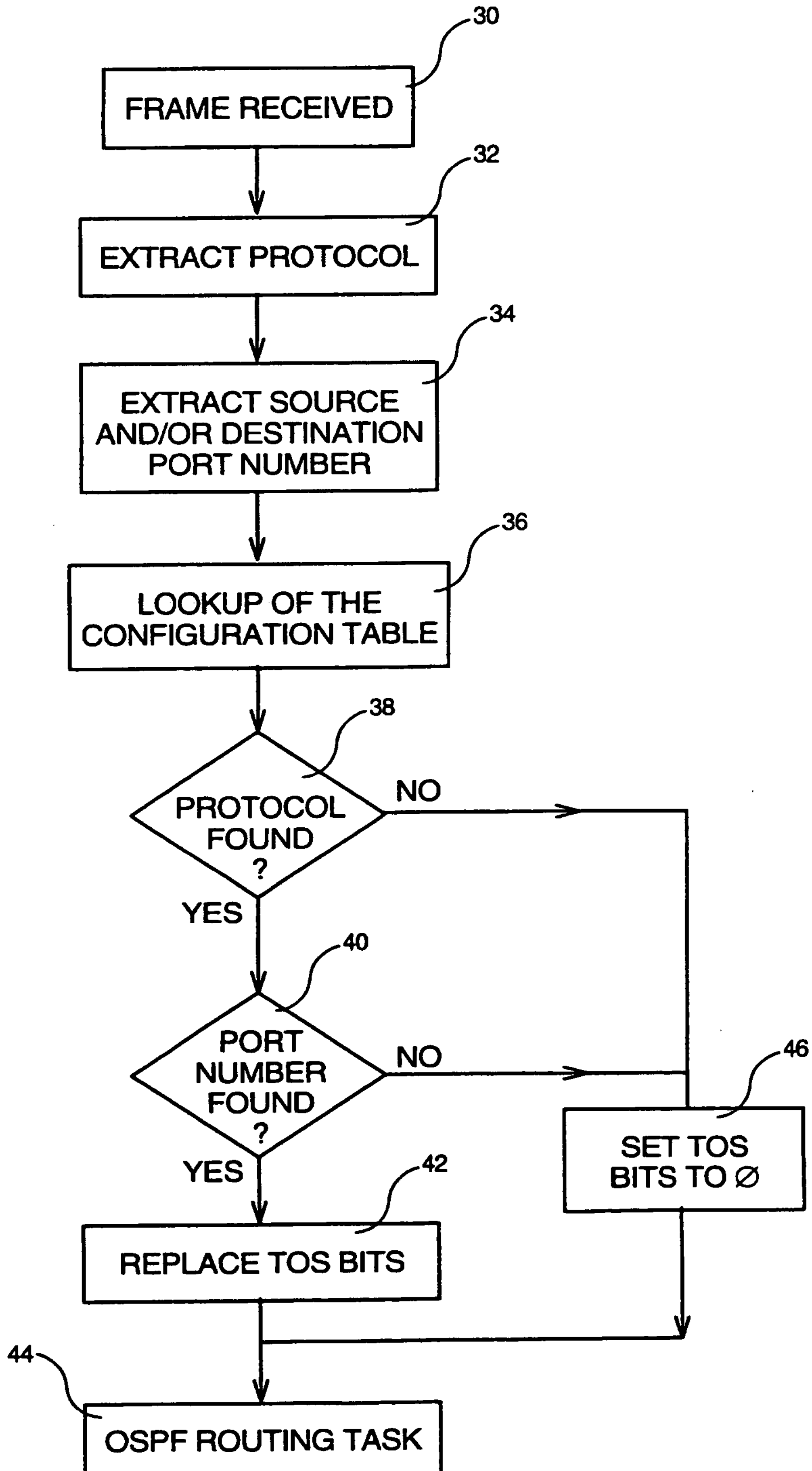

FIG. 5

FR 9 00 008
Benayoun et al
4/4

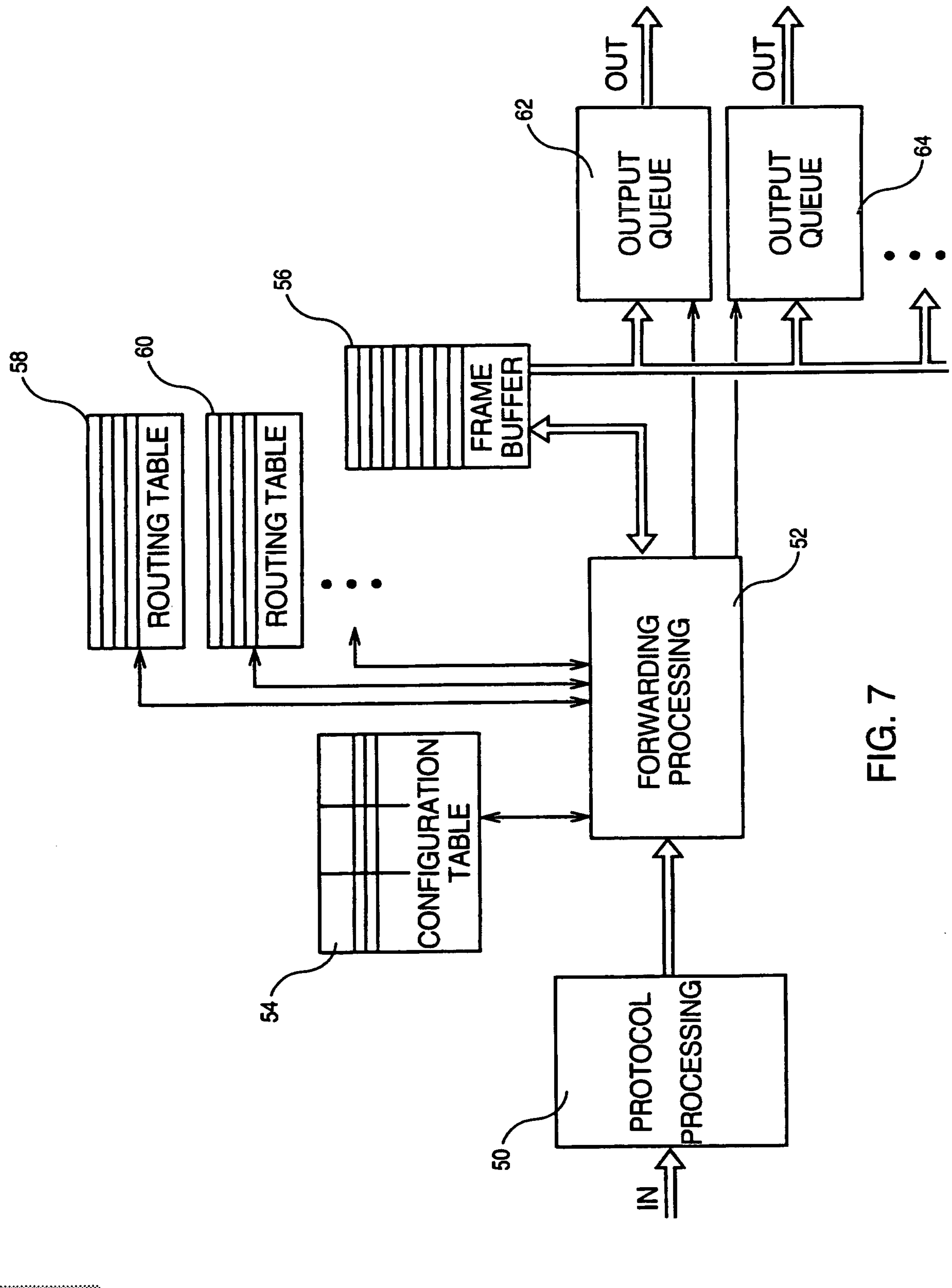

FIG. 7